Food and Agriculture
Organization of the
United Nations

Forest monitoring good practices

Mapping tree crop commodities

SHOUT
SHARE
IT

Forest monitoring good practices

Mapping tree crop commodities

Authors

Rémi d'Annunzio, Jacob Abramowitz (The University of Alabama in Huntsville), Fernand Bale (BNETD), Nicholas Clinton (Google), Adeline Dontenville (EFI), Karen Dyson (SIG), Luca Foresta (Satelligence), Juliana Freitas Beyer (Thünen Institute of Forestry), Lavinia Gasperini (International Coffee Organization), Patrick Houdry (Airbus), Neha Hunka (ESA), Örjan Jonsson, Erik Lindquist, Melvin Lippe (Thünen Institute of Forestry), Baptiste Monnier (Sucden Coffee), Fatima Mushtaq, Kasimir Orlowski (FINCONS S.P.A), Pierrick Poulenas (Picterra), Pierrick Rambaud (LDC), Felix Rembold (EC-JRC), Louis Reymondin (Alliance Bioversity-CIAT), Carlos Riano (EFI), Vivian Ribeiro (Meridia), Pascal Ripplinger (GIZ), David Saah (SIG), Laurent Sagarra (JDE Peet's), Christophe Sannier (GAF), Sam Schechtman (JDE Peet's), Aurélie Shapiro, Alicia Sullivan (Google), Karis Tenneson (SIG), Maria Pilar Valbuena Perez, Thibaud Vantalon (Alliance Bioversity-CIAT), Andreas Vollrath and Benjamin Wielgosz (WRI).

Food and Agriculture Organization of the United Nations

Rome, 2026

Required citation:

d’Annunzio, R., Abramowitz, J., Bale, F., Clinton, N., Dontenville, A., Dyson, K., Foresta, L., Freitas Beyer, J., Gasperini, L., Houdry, P., Hunka, N., Jonsson, Ö., Lindquist, E., Lippe, M., Monnier, B., Mushtaq, F., Orlowski, K., Poulenas, P., Rambaud, P., Rembold, F., Reymondin, L., Riano, C., Ribeiro, V., Ripplinger, P., Saah, D., Sagarra, L., Sannier, C., Schechtman, S., Shapiro, A., Sullivan, A., Tenneson, K., Valbuena Perez, M.P., Vantalon, T., Vollrath, A. & Wielgosz, B. 2026. ***Forest monitoring good practices: mapping tree crop commodities***. Rome, FAO. https://doi.org/10.4060/cd9622en

The designations employed and the presentation of material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations (FAO) concerning the legal or development status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO in preference to others of a similar nature that are not mentioned.

ISBN 978-92-5-140699-1

© FAO, 2026

Some rights reserved. This work is made available under the Creative Commons Attribution- 4.0 International licence (CC BY 4.0: https://creativecommons.org/licenses/by/4.0/legalcode.en).

Under the terms of this licence, this work may be copied, redistributed and adapted, provided that the work is appropriately cited. In any use of this work, there should be no suggestion that FAO endorses any specific organization, products or services. The use of the FAO logo is not permitted. If a translation or adaptation of this work is created, it must include the following disclaimer along with the required citation: “This translation [or adaptation] was not created by the Food and Agriculture Organization of the United Nations (FAO). FAO is not responsible for the content or accuracy of this translation [or adaptation]. The original [Language] edition shall be the authoritative edition.”

Any dispute arising under this licence that cannot be settled amicably shall be referred to arbitration in accordance with the Arbitration Rules of the United Nations Commission on International Trade Law (UNCITRAL). The parties shall be bound by any arbitration award rendered as a result of such arbitration as the final adjudication of such a dispute.

Third-party materials. This Creative Commons licence CC BY 4.0 does not apply to non-FAO copyright materials included in this publication. Users wishing to reuse material from this work that is attributed to a third party, such as tables, figures or images, are responsible for determining whether permission is needed for that reuse and for obtaining permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

FAO photographs. FAO photographs that may appear in this work are not subject to the above-mentioned Creative Commons licence. Queries for the use of any FAO photographs should be submitted to: photo-library@fao.org.

Sales, rights and licensing. FAO information products are available on the FAO website (www.fao.org/publications) and print copies can be purchased through the distributors listed there. For general enquiries about FAO publications please contact: publications@fao.org. Queries regarding rights and licensing of publications should be submitted to: copyright@fao.org.

Cover and back cover photographs: © FAO/Pilar Valbuena

Contents

Contents ... iii

Acknowledgements ... iv

Abbreviations ... v

Executive summary ... vi

Introduction and context ... 1

Definitions ... 2

Rationale ... 4

- Structural constraints ... 4
- Typology considerations ... 5

Approaches to commodity mapping ... 6

- Classification system ... 7
- Earth observation inputs ... 7
- Machine-learning approaches ... 10
- Reference data ... 11
- Postprocessing ... 13
- Accuracy assessment ... 14
- Convergence of evidence ... 15

Data sharing ... 16

- Availability of end results ... 16
- Availability of metadata ... 16
- Availability of reference data ... 17
- Towards a pool of reference data ... 17

Conclusion ... 20

References ... 21

Data protection laws ... 24

Existing initiatives ... 25

Acknowledgements

The forest monitoring good practices on mapping tree crop commodities were made possible thanks to the contributions of many dedicated people and institutions that willingly provided time and specialist expertise during the formulation and review process.

The preparation of the good practices was coordinated by Rémi d'Annunzio (FAO). The authors were Jacob Abramowitz (The University of Alabama in Huntsville), Fernand Bale (BNETD), Nicholas Clinton (Google), Adeline Dontenville (EFI), Karen Dyson (SIG), Luca Foresta (Satelligence), Juliana Freitas Beyer (Thünen Institute of Forestry), Lavinia Gasperini (International Coffee Organization), Patrick Houdry (Airbus), Neha Hunka (ESA), Örjan Jonsson (FAO), Erik Lindquist (FAO), Melvin Lippe (Thünen Institute of Forestry), Baptiste Monnier (Sucden Coffee), Fatima Mushtaq (FAO), Kasimir Orlowski (FINCONS S.P.A), Pierrick Poulenas (Picterra), Pierrick Rambaud (LDC), Felix Rembold (EC-JRC), Louis Reymondin (Alliance Bioversity-CIAT), Carlos Riano (EFI), Vivian Ribeiro (Meridia), Pascal Ripplinger (GIZ), David Saah (SIG), Laurent Sagarra (JDE Peet's), Christophe Sannier (GAF), Sam Schechtman (JDE Peet's), Aurélie Shapiro (FAO), Alicia Sullivan (Google), Karis Tenneson (SIG), Maria Pilar Valbuena Perez (FAO), Thibaud Vantalon (Alliance Bioversity-CIAT), Andreas Vollrath (FAO) and Benjamin Wielgosz (WRI).

Copyediting was completed by Leslie Saffrey, with graphic design by Anastasiya Lapova. This paper was produced by FAO under its Accelerating innovative monitoring for forests programme (AIM4Forests) and Accelerating innovative monitoring for forests and commodities initiative (AIM4Commodities), thanks to finance from the German Agency for International Cooperation (GIZ) and the Department for Energy Security and Net Zero of the United Kingdom (DESNZ).

Abbreviations

AI	artificial intelligence
CC-BY-4.0	Creative Commons Attribution 4.0 International License
CNA	competent national authority
CoE	convergence of evidence
DIASCA	Digital Integration of Agricultural Supply Chains Alliance
DPI	digital public infrastructure
EC	European Commission
EFI	European Forest Institute
EO	Earth observation
EU	European Union
EUDR	European Union Deforestation Regulation
FAO	Food and Agriculture Organization of the United Nations
FPIC	free, prior and informed consent
GHG	greenhouse gas
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
IP	intellectual property
ISO	International Organization for Standardization
JRC	Joint Research Centre
LCML	land cover meta language
LiDAR	light detection and ranging
PII	personally identifiable information
SASI	Sustainable Agricultural Supply Chains Initiative
TF	Technical Facility on Deforestation-Free Value Chains

Executive summary

This document sets out good practices for mapping tree crop commodities – such as cocoa, coffee, rubber, palm oil and agroforestry systems – using Earth observation and geospatial methods. It clarifies key definitions (*land cover, land use, deforestation, mapping, agroforestry*) and explains why distinguishing between forests and tree-based agricultural land uses is critical for addressing commodity-driven deforestation and meeting rising regulatory and market expectations for deforestation-free supply chains. While numerous public and private initiatives now produce commodity maps, misalignment in definitions, methods, data quality and transparency undermines trust and comparability. The good practices therefore aim to promote consistency, credibility and interoperability across regions and commodities, while remaining non-binding and flexible to national circumstances.

The document also outlines the major technical and governance challenges in mapping commodities with tree cover. These include heterogeneous and dynamic production systems, limitations of remote sensing data – particularly in shaded agroforestry systems, lack of consistent typologies, limited access to high-quality reference data and reluctance to share data due to commercial, legal or privacy concerns.

To address these issues, the document emphasizes good practices in data collection, modelling, validation, metadata and data sharing, including hybrid approaches that combine remote sensing with field data, locally grounded accuracy assessments, convergence-of-evidence frameworks, and open, well-documented outputs. They further highlight the importance of transparent data governance, appropriate licensing, informed consent and compliance with data protection laws, to ensure that commodity maps support regulatory compliance, responsible sourcing, inclusion of smallholders and informed decision-making across global supply chains.

Introduction and context

The expansion of agricultural land linked to the production of various commodities remains the main driver of deforestation globally (Branthomme ***et al.***, 2023). Tree crop commodities contribute to a non-negligible part of the expansion of agricultural land and forest removal. As new regulatory mechanisms and voluntary commitments for more sustainable sourcing practices are on the rise to curb deforestation linked to supply chains, monitoring solutions are needed to identify where they are grown. Those solutions feed on robust spatial information that needs to be comparable across countries and commodities.

Several public and private initiatives are already generating datasets showing the presence or absence of those commodities at key points in time, making the resulting commodity maps available for the global community. This momentum carries a responsibility of the different initiatives to steer towards alignment in order to bring trust in the geospatial data feeding monitoring solutions.

The goal of these good practices is to promote consistency, transparency and credibility in the mapping of tree crop commodities across regions and supply chains.

Mapping tree crop commodities is inherently complex because remote sensing sensors capture biophysical land-cover characteristics rather than land use. Tree crop commodities systems often resemble natural forests in Earth observation (EO) data, complicating the differentiation between forest and agriculture. The land-use disambiguation of commodities with tree cover versus forests is the key focus of the current good practices.

The main challenges associated with this disambiguation can be grouped into interrelated domains: landscape and vegetation complexity, technical constraints of remote sensing systems, reference data limitations, and cost and scalability constraints.

Additional challenges include inconsistent definitions across countries, jurisdictions, limited access to ground reference and training datasets, capacity constraints in some countries, and concerns related to data sensitivity and ownership.

© FAO/Pilar Valbuena

Overcoming these systemic limitations requires not only technological innovation but also methodological and institutional alignment. Current mapping initiatives often apply different classification schemas, validation protocols and spatial resolutions tailored to specific commodity systems or geographies. While these approaches serve local needs, they fragment the global landscape of agricultural mapping and hinder comparability across datasets.

Establishing a shared foundation of definitions, accuracy metrics and documentation standards would improve transparency and reproducibility, allowing decision-makers to build trust in the data that underpin deforestation-free monitoring frameworks. Such convergence does not imply uniformity but rather consistency, promoting alignment around good practices while allowing flexibility for distinct contexts and capacities.

These good practices provide a common reference for governments, private sector actors, civil society and researchers to align methodologies while allowing flexibility for national contexts. They support responsible decision-making, reduce duplication of efforts, and foster trust among stakeholders by clarifying good practices without imposing binding requirements.

In addition to supporting regulatory compliance on zero-deforestation, common methods for mapping commodity presence and change over time enable consistency across different processes that require quantification of deforestation such as greenhouse gas (GHG) protocol baselines and reporting with sustainability targets.

In general, the document provide a methodological “safe harbour” that reduces compliance friction and can contribute to data being consistently accepted across entire value chains, from smallholder producers and suppliers to operators and regulators.

Definitions

A standardized classification framework to describe land cover and land use helps ensure that these concepts are consistently applied across mapping initiatives and monitoring systems. This is particularly important in the context of the mapping of tree crop commodities, where key terms such as *land cover, land use, deforestation, mapping*, and *agroforestry* are often interpreted differently across jurisdictions and datasets.

For this reason, we first clarify definitions and terms used in the document.

AGROFORESTRY

The Food and Agriculture Organization of the United Nations (FAO) definition of agroforestry is the collective term for land management systems where woody perennials (trees, shrubs, palms, bamboos, etc.) are deliberately integrated with agricultural crops, animals or both in some form of spatial arrangement or temporal sequence. Agroforestry is a subcategory of “other land with tree cover”.

DEFORESTATION

The Food and Agriculture Organization of the United Nations (FAO) defines deforestation as the conversion of forest to other land use independently whether human induced or not. As such, deforestation does not necessarily imply loss of tree cover (land cover) but it does imply conversion of forests into another land use.

FOREST

The Food and Agriculture Organization of the United Nations (FAO, 2023) land use–based definition of forest is adopted here – land spanning more than 0.5 hectares with trees higher than 5 metres and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use. Of note, this excludes tree stands in agricultural production systems, such as fruit tree plantations, oil palm plantations, olive orchards and agroforestry systems when crops are grown under tree cover.

LAND COVER

"Land cover is the observed (bio)physical cover on the Earth's surface" (Di Gregorio & Jansen, 2000). "Tree cover" and "grassland" are land-cover concepts.

LAND USE

"Land use is characterized by the arrangements, activities and inputs people undertake in a certain land-cover type to produce, change or maintain it. Definition of land use in this way establishes a direct link between land cover and the actions of people in their environment" (Di Gregorio & Jansen, 2000). "Forest" and"pasture" are land-use concepts.

MAPPING

Mapping refers to the systematic spatially explicit representation of land cover, land use and agricultural commodities using geospatial data and analytical methods. In this context, mapping includes data acquisition, processing, classification, validation and sharing of spatial information related to tree crop commodities.

We are specifically focusing in this paper on land-use mapping of commodities, i.e. generating a geospatial map of the presence or probability of presence of the commodities, at the pixel level.

In particular, we do not mean mapping and delineating the boundaries of farms or fields, even though those datasets may serve as intermediaries to achieve the end results.

TREE COVER

Tree cover is defined by the presence of woody vegetation of a minimum height, canopy cover or both. This is a structural and physical definition that may include "planted forests, tree crops (such as oil palm or apple orchards) as well as managed and unmanaged natural forests" (WRI, 2023). Tree canopy cover is a component of the forest definition.

TREE CROP COMMODITIES

In this paper, we focus on commodities growing on trees or shrubs like cocoa, coffee, oil palm, rubber and cashew. These commodities grow in varying types of systems, ranging from monoculture full-sun plantations to mixed-cropping systems and dense agroforestry with a closed shade-tree canopy.

Rationale

Structural constraints

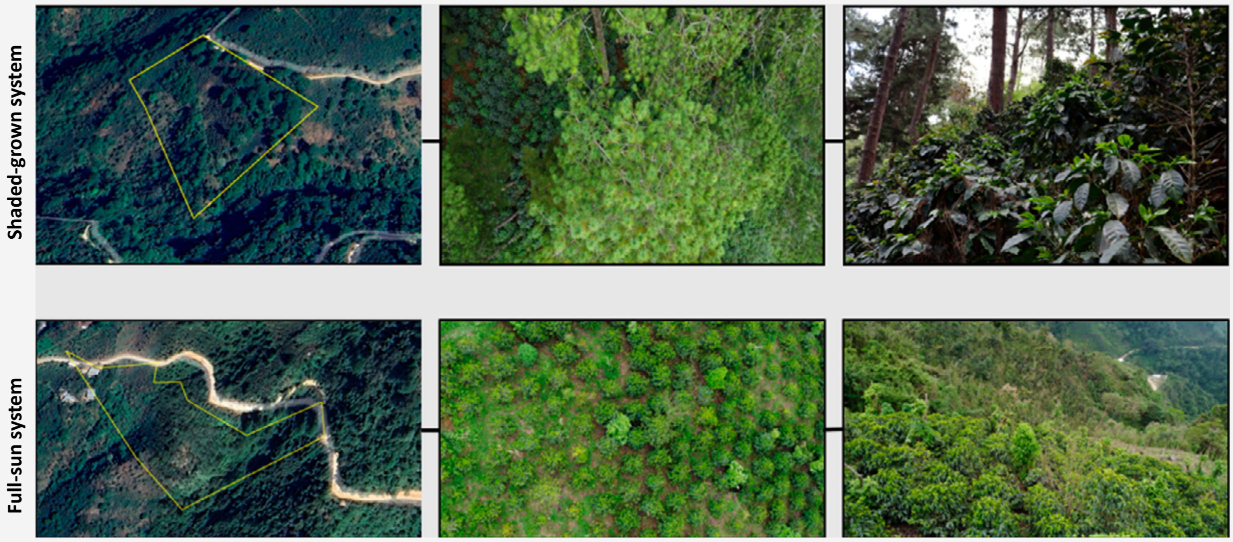
Tree crop commodities systems are implemented across a wide range of canopy structures, management intensities and ecological contexts. Coffee, for example, may be cultivated under dense multistrata shade systems, commercial polycultures or simplified shaded monocultures. Cocoa and rubber are frequently planted within existing natural forest canopies or interspersed with fallow land that closely resembles regenerating vegetation in structure and composition.

These systems often occur within dynamic landscape mosaics where agroforestry plots, secondary vegetation and natural forest coexist. Successional processes and management practices, including pruning, shade thinning, temporary shade establishment and replanting, further increase variability over time.

Landscape configuration plays a critical role in the separability of land-cover classes. In highly fragmented settings, agricultural plots and forested areas can form fine-scale mosaics with irregular boundaries. This increases the extent of transition zones between adjacent land-cover types, reducing class separability and contributing to classification uncertainty in mapping applications. Moreover, very small, elongated or irregularly shaped patches increase boundary density and edge effects, complicating consistent delineation.

Agroforestry systems exemplify these challenges. The mixture of cultivated and naturally occurring species, differences in plant structure and density, and the presence of shade trees create high structural similarity with forest. A systematic literature review on coffee mapping (Escobar-López *et al*., 2024) has shown that in the case of agroforestry, classification accuracy decreases as canopy density increases, especially when tree cover exceeds roughly 45 percent, leading to confusion with secondary vegetation or other agroforestry systems. In many agroforestry regions, transitions between cultivated systems and natural forest are gradual rather than sharply bounded. These diffuse boundaries reflect ecological and management gradients rather than discrete land-use changes, further complicating consistent delineation.

Figure 1. *Two examples of coffee agroforestry systems (shade grown and full sun) from the southern part of Chiapas province, Mexico, showing the complexity for Earth observation systems to map heterogeneous land uses. Photos taken in June 2025 using DJI Mini 4 Pro at an altitude of 50 m and Sony EOS 250D for ground-level; © Melvin Lippe/ Thünen Institute. Google Earth imagery: © 2025 Airbus. Image date: January 8, 2024. Photos: © Lippe Melvis/ Thünen Institute.*



Satellite imagery- *good canopy detail, very limited subcanopy visibility*

Drone imagery- *full canopy detail, partial subcanopy visibility*

Ground-level photo - *full subcanopy detail*

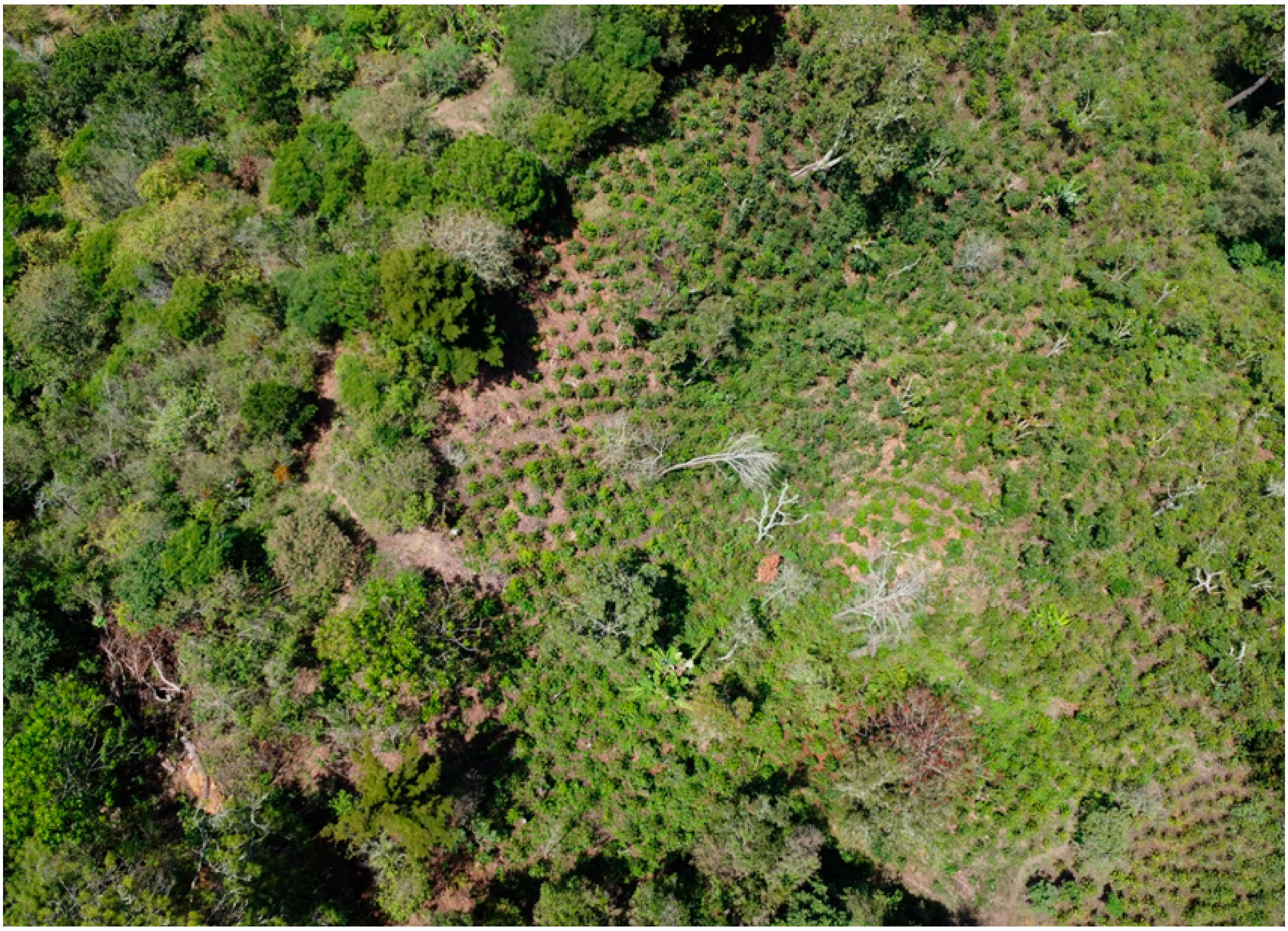
In highly shaded systems, differentiation between agroforestry and natural forest may be impossible due to the structural similarity between cultivated and naturally occurring vegetation. In dense multistrata systems, subcanopy crops may remain physically obscured by upper canopy layers, representing inherent structural visibility limits. At the same time, structural complexity spans a continuum, extending to young or sparsely planted stands with incomplete canopy closure, where heterogeneous canopy structure and exposed ground and herbaceous vegetation further increase within-class variability.

Typology considerations

The absence of a clear typology to designate agroforestry types in scientific literature applied to remote sensing–based products further confuses the interpretation of the landscape. Existing studies illustrate significant inconsistencies in how agroforestry systems are defined across different countries and agroecological regions. Mapping efforts often aggregate multiple crops into broad "industrial crops" or "agriculture" categories, obscuring important distinctions among shade systems and traditional or hybrid tree crop varieties.

Only a few efforts explicitly differentiate levels of shade or intercropping. In many agroforestry regions there are no clear plot boundaries but rather gradual transitions among forest types which makes it hard even for those most advanced machine learning and artificial intelligence (AI) foundation models to clearly distinguish biophysical boundaries between where natural forests end and agroforestry begins.

Figure 2. *Drone imagery of two adjunct coffee farming systems in the northern part of Oaxaca province, Mexico. The left part shows a more traditional agroforestry system with a high density of shade trees while the right part refers to a monocultural plot with almost no shade trees. Photo taken in May 2025 using DJI Mini 4 Pro at an altitude of 50 m; © Melvin Lippe/Thünen Institute.*



Approaches to commodity mapping

© FAO/Pilar Valbuena

Despite those structural challenges, literature reviews and key publications on mapping palm (Descals *et al.*, 2024), cocoa (Kalischek *et al.*, 2023; Otekunrin, 2025), coffee (Hunt *et al.*, 2020; Maskell *et al.*, 2021; Escobar *et al.*, 2024; Parreiras *et al.*, 2025) and rubber (Yue and Xiao , 2025) show that several hundred articles on these topics have been published in the past five years, with recent years showing a marked acceleration in research output due to technological advances such as high-resolution (both spatial and temporal) imagery and AI.

A non-exhaustive selection of meta reviews on the topic can be found in the "Mapping commodities" section of the References.

In general, mapping a commodity through EO data requires:

1. classification system
2. EO inputs
3. inference models
4. reference data for calibration and validation
5. postprocessing protocols
6. accuracy assessment standards

The following sections propose some good practices on each of these six components.

Classification system

The International Organization for Standardization (ISO) 19144 standard series (land cover meta language; LCML) provides a standardized framework for characterizing land cover and land use. Its application helps ensure consistent interpretation of key concepts such as forests, plantations, and agroforestry systems across countries and datasets.

By supporting a structured description of vegetation characteristics and land-use functions, the standard enables clearer differentiation between natural forests and tree-based agricultural systems such as oil palm plantations or cocoa agroforestry. This is particularly important for monitoring commodity-driven deforestation and ensuring consistency in mapping approaches.

The use of standardized classification systems also improves the interoperability of datasets, facilitates the integration of field observations with satellite data, and supports applications such as supply chain transparency, deforestation monitoring, carbon accounting and climate reporting. It is good practice, when designing a land-cover or land-use map, to define the classes using LCML.

Earth observation inputs

Several types of EO inputs can be used individually or in combination to generate maps: they typically are satellite-based optical imagery, radar and light detection and ranging (LiDAR) data (JAXA, 2003).

Optical satellite images capture reflected sunlight across different wavelengths, allowing vegetation, water, soil and urban areas to be distinguished based on their spectral signatures. They come with different spatial, temporal and spectral resolution characteristics.

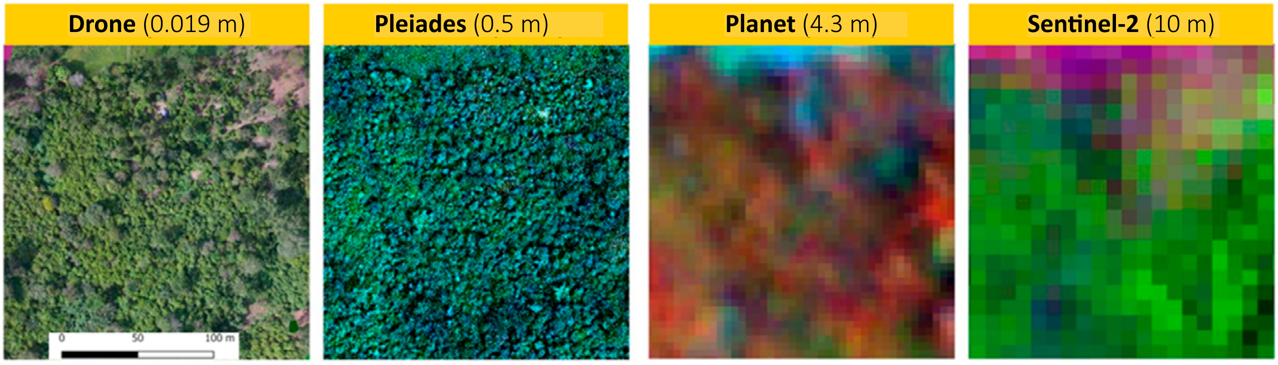
SPATIAL RESOLUTION

Medium-resolution imagery (30 m to 10 m) often aggregates multiple land uses within a single pixel, particularly in smallholder contexts where plots are small and intercropped. Increasing spatial resolution improves the detectability of individual tree crowns and small production units, thereby increasing the potential for differentiating certain tree commodities.

However, very high-resolution imagery (e.g. submetre or centimetre range) changes the scale at which landscapes are interpreted. Features that are not visible at lower resolution, such as small canopy gaps, shade-tree variability, or fine-scale structural differences, become detectable. This increased detail does not necessarily simplify classification but instead requires more precise and scale-consistent class definitions.

© FAO/Pilar Valbuena

Figure 3. *Comparison of spatial resolution across remote sensing platforms over a cocoa agroforestry plantation. From left to right: drone imagery (0.019 m), Pléiades (0.5 m), PlanetScope (4.3 m) and Sentinel-2 (10 m). The images demonstrate how increasing pixel size affects the detectability of agroforestry features, ranging from individual cocoa trees and shade-tree crowns in very high-resolution drone data to broader canopy patterns at medium resolution. Image sources: Drone: GIZ; Pléiades: © CNES (2020), distribution Airbus Defence and Space; Planet: © 2020 Planet Labs PBC. Figure prepared by the European Commission Joint Research Centre (EC-JRC).*



For example, gaps between cocoa stands that are indistinguishable at medium resolution may become clearly visible at very high resolution. This raises definitional questions: At what gap size does an opening remain part of a cocoa-planted area, and at what point does it constitute a separate non-cocoa land cover? Without clearly defined minimum mapping units and classification rules, increasing spatial resolution can lead to inconsistent labelling.

Higher-resolution imagery shifts part of the challenge from detecting the presence of a commodity to defining how it should be represented. As finer spatial detail reveals internal heterogeneity, such as canopy gaps, shade trees, and structural variability within plots, classification performance increasingly depends on consistent class definitions, clearly defined minimum mapping units, and sufficiently detailed reference data. While increased spatial detail can improve mapping in some contexts, it does not overcome structural visibility constraints in dense, multistrata systems where subcanopy crops remain physically obscured.

SPECTRAL RESOLUTION

Optical remote sensing relies on reflectance properties to infer land cover. However, dense agroforestry systems and natural forests often share similar spectral signatures. The absence of rich multispectral information in some sensors further limits detailed land-cover interpretation.

Hyperspectral sensors provide detailed, contiguous spectral information that can improve discrimination between agroforestry systems and natural forests by capturing subtle differences in vegetation composition and condition. However, their limited spatial coverage, high data volume, restricted availability and inability to overcome structural occlusion in dense multistrata systems currently constrain large-scale operational use.

CLOUD COVER AND NON-OPTICAL SENSORS

In addition, optical sensors depend on clear skies and daylight, which can limit data availability in cloudy regions. In many commodity-producing regions, persistent cloud cover severely constrains the availability of usable optical imagery. This is particularly problematic for high-resolution imagery with low or irregular acquisition frequency. As a result, high-resolution datasets capable of distinguishing subcanopy land cover may be outdated or entirely unavailable.

Radar sensors overcome this limitation and can penetrate cloud cover and partially alleviate data gaps in tropical regions by actively sending microwave signals and measuring their return from the Earth's surface. Because radar can penetrate clouds and operate day or night, it provides consistent observations of surface structure, moisture and roughness.

However, publicly available spaceborne radar datasets typically provide comparatively coarse spatial resolution, and the complexity of backscatter signals can limit interpretability for detailed land-cover discrimination. Higher-resolution radar systems exist but are not consistently available at broad spatial scales. Airborne or spaceborne LiDAR datasets are well suited for characterizing vertical vegetation structure, yet their spatial coverage remains limited.

COST AND SCALABILITY CONSTRAINTS

Even where technical improvements are possible, cost presents a major barrier.

The acquisition and preprocessing of commercial submetre optical imagery at scale is expensive. Airborne and spaceborne LiDAR datasets are constrained by very high expense and limited spatial coverage. Collecting detailed and locally specific ground reference data is also resource-intensive and often limited by capacity constraints.

Some technical challenges, such as improving spatial detail or integrating radar data to address cloud cover, can be partially mitigated with higher investment. However, other constraints, particularly those related to dense canopy occlusion and structural similarity between agroforestry and natural forest, persist regardless of cost.

In parallel to these technical and financial barriers, institutional and data governance challenges strongly influence the feasibility of scalable mapping efforts. The collection, curation and distribution of high-quality geospatial data often hinge on fragmented policies, proprietary datasets and uneven national capacities. Licensing conditions, data ownership concerns and privacy regulations can limit the accessibility and reuse of existing information, even where technical capabilities exist.

The policy, ownership and capacity fragmentation discourages interoperability between datasets and reinforces disparities between well-resourced and data-poor regions. As a result, the obstacles to achieving consistent, global-scale mapping extend beyond sensor performance and cost: They also stem from institutional fragmentation and the absence of harmonized standards for data stewardship and sharing.

COMPLEMENTARITY BETWEEN ACTIVE AND PASSIVE SENSORS

When optical and radar data are combined, they provide complementary information: Optical data reveals colour and vegetation properties, while radar highlights surface texture and physical structure. This integration improves classification accuracy and reliability, enabling researchers and decision-makers to monitor forests, agriculture, urban expansion and environmental changes effectively through detailed and up-to-date land-cover maps.

Emergent technologies such as Alpha Earth Embeddings and other geofoundation models hold promise for their convenient synthesis of information from a variety of sensors and spectral regions, but the lack of interpretability requires very accurate ground reference and a machine-learning model to make predictions.

Machine-learning approaches

© FAO/Sarah Pechevis

The production of commodity maps through conventional pixel classification has a track record of success in both public and commercial sectors, particularly where the crops are of uniform cover such as monoculture row-crops. However, when such commodities are grown under mixed cover such as in silvo-pasture, agroforestry or multicropping systems, those conventional models can easily confuse managed production areas with natural forest cover or mixed canopy cover of non-forest ecosystems.

Many of the highest accuracy crop-mapping initiatives have leveraged primary ground data in combination with deep-learning or AI foundation models, applied to a regional or commodity specific objective such as soy, palm or cocoa.

As a complement or alternative, probability maps derived from regressions rather than classification approaches can provide meaningful area-level estimation of production areas or function as a filter within analyses, allowing for explicit choice or statement of the uncertainty of results. They may be useful in guiding validation exercises by speeding the production of relevant validation points by annotation teams and also act as a secondary indication of the spatial variability, complementing validation results based on a sample representative of a given area and not necessarily applicable at pixel level.

In addition, the availability of probabilities for both commodity and forest cover allow the user to set their own threshold that would be applicable over a certain area to produce their own commodity mapping layers maps.

Reference data

Accurate model development depends on detailed, representative and accessible reference data. However, limitations arise not only from technical constraints but also from institutional, economic and governance factors that restrict reference data availability and quality.

AVAILABILITY AND ACCESSIBILITY

Reference datasets are frequently not openly shared due to proprietary interests, competitive concerns, licensing restrictions or privacy considerations related to smallholder farm boundaries. In some commodity sectors, agricultural statistics and production data carry commercial sensitivity. Furthermore, commodity maps may challenge the reliability of official registries by revealing unregistered production areas or cultivation within protected zones, which can discourage data transparency. In addition, institutional and technical capacity constraints in some regions limit systematic data collection, maintenance and curation. As a result, comprehensive and up-to-date reference datasets are often unavailable.

CLASS GRANULARITY AND TYPOLOGICAL CLARITY

The absence of a clear and widely accepted typology for agroforestry and tree crop systems further complicates reference data design. Definitions of agroforestry vary substantially across countries and agroecological regions, and mapping efforts often aggregate diverse systems into broad categories such as “industrial crops” or “agriculture”. Such aggregation obscures important distinctions among shade levels, intercropping practices, and traditional or hybrid cultivation systems. When these distinctions are not reflected in the reference data, models are constrained to learning commodity presence–absence patterns and cannot differentiate classification outputs according to specific cultivation practices or management intensity.

REPRESENTATION OF NON-TARGET CLASSES

Reference datasets often emphasize the target commodity while providing limited or poorly defined samples of adjacent or structurally similar non-target classes, such as secondary vegetation, fallow land, degraded forest or mixed agroforestry systems. Insufficient representation of these classes restricts the model’s ability to learn meaningful decision boundaries and increases confusion in heterogeneous landscapes. Without balanced and well-characterized non-target references, classification models may exhibit elevated commission errors and reduced generalization performance.

ANNOTATION PRECISION AND SCALE CONSISTENCY

Even when commodity labels are available, spatial delineation is often coarse or generalized. As spatial resolution increases and internal heterogeneity within production systems becomes visible, precise annotation detail and clearly defined minimum mapping units become increasingly important. Reference datasets that lack scale-consistent delineation standards may therefore be inadequate for training or validating high-resolution classifications. Without such scale-consistent reference data, improvements in sensor resolution do not necessarily translate into improved classification performance.

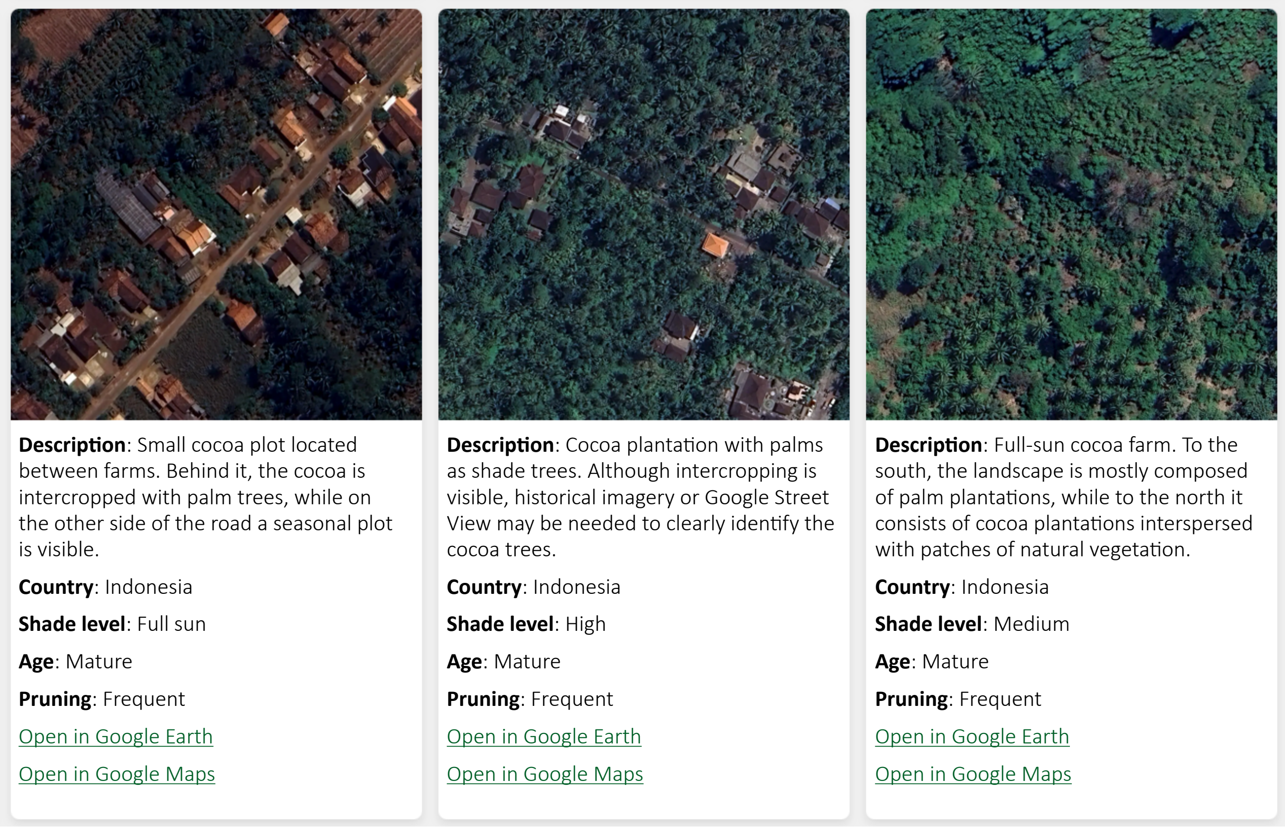
Together, these limitations constrain not only model performance but also the interpretability and comparability of remote sensing–based commodity mapping products.

GOOD PRACTICES FOR REFERENCE DATA COLLECTION

Because of the wide variety of configurations for agrocommodity systems that incorporate natural tree cover or arboreal crops, a lot of reference data is required to capture the range of possible conditions. In addition to being of sufficient density in feature space, the data need to be well distributed in space (and possibly time) in order to generalize to larger areas and for several time periods.

Both visually interpreted aerial imagery and field information can be used as complementary sources of training and validation data for the models (Orlowski et al., 2026). In addition, fieldwork for complex and less investigated systems is critical for training skilled labellers and increasing data quality.

Figure 4. *Examples of interpretation guide for different cocoa growing systems in Indonesia. Data extracted from the Commodity Recognition Book. © Forest Data Partnership*



Description: Small cocoa plot located between farms. Behind it, the cocoa is intercropped with palm trees, while on the other side of the road a seasonal plot is visible.

Country: Indonesia

Shade level: Full sun

Age: Mature

Pruning: Frequent

Open in Google Earth

Open in Google Maps

Description: Cocoa plantation with palms as shade trees. Although intercropping is visible, historical imagery or Google Street View may be needed to clearly identify the cocoa trees.

Country: Indonesia

Shade level: High

Age: Mature

Pruning: Frequent

Open in Google Earth

Open in Google Maps

Description: Full-sun cocoa farm. To the south, the landscape is mostly composed of palm plantations, while to the north it consists of cocoa plantations interspersed with patches of natural vegetation.

Country: Indonesia

Shade level: Medium

Age: Mature

Pruning: Frequent

Open in Google Earth

Open in Google Maps

Interpretation guides, adapted to the local context, are necessary so that labelling technicians can ensure consistency of the interpretation. Equally important is the development of a strong network of well-trained local technicians who understand production systems in the context of local landscapes. Such networks improve labelling accuracy and help ensure that models remain robust and contextually relevant at the local scale.

A non-exhaustive selection of publications on the topic can be found in the “Good practices for data collection” section of the References.

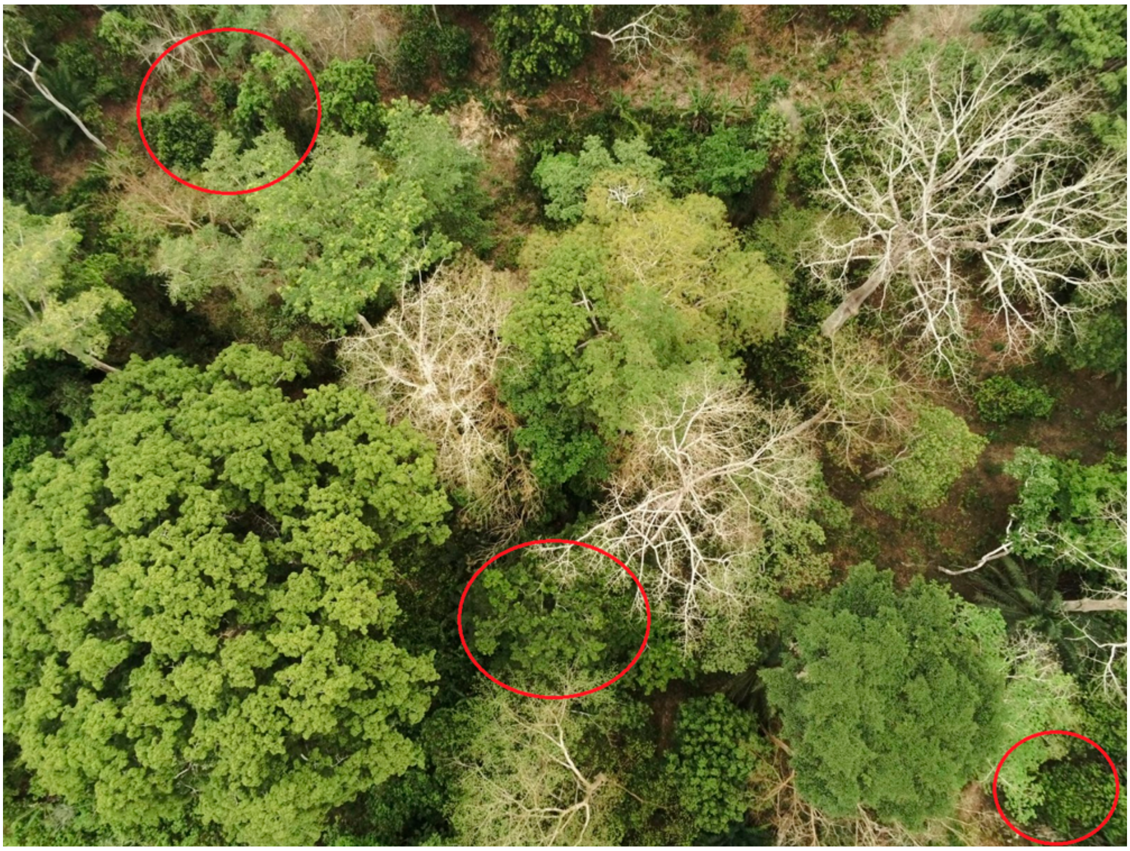
Postprocessing

Advanced approaches using AI may provide solutions to discern natural forests from agroforestry systems, but these require a great amount of widely distributed and high-quality field data.

High-resolution optical imagery is often used to create reference data over large areas, especially when field data collection may be difficult or expensive to collect. But the limitations of optical imagery need to be clear and understood, in particular for closed canopy systems such as cocoa and coffee, making ground-truthing data more important and relevant than even high spatial resolution acquisitions.

In complex landscapes such as in Cameroon, highly shaded cocoa plantations reach these limits and affect the collection of data through very high-resolution image annotation. While the identification of certain land-cover classes, such as palm, rubber and grassland, is possible to determine visually, the differentiation between natural forests and agroforestry systems is in most cases impossible – even at highest resolution. For those reasons, the current capabilities of remote sensing techniques and data for full-shaded agroforestry systems perform poorly: "we can't see well below the canopy".

Figure 5. *Example of a highly varied, multistory agroforestry cocoa plantation in Cameroon, imaged from a drone at a height of approximately 50 m. While some cocoa trees can be seen through canopy gaps (red circles), they are mostly hidden by taller trees. © FAO/Aurélie Shapiro*



In those cases, it may be necessary to directly complement model-derived products with manual inclusion of field delineated data (farm boundaries). The information on field boundaries can be merged into the commodity map but geolocation information or field boundaries are usually considered personally identifiable information (PII) by data protection laws. The rasterization process of the field boundaries provides a natural anonymization step that supports compliance with those laws.

Accuracy assessment

Model-derived geospatial data products (e.g. remote sensing–based land-cover maps) must be statistically validated to determine their accuracy level (Stehman &Foody, 2019). The validation is performed through a probability sampling design, using independent reference data to compare against the map. A non-exhaustive selection of publications can be found in the “Accuracy assessment of land cover maps” section of the References.

Validation should explicitly quantify confusion between policy-relevant classes and not only overall accuracy. When assessing exposure to non-compliance risks related to a given commodity expansion into forests, the minimum validation legend should include at least three classes: the commodity, forest and other lands. A confusion matrix should then be computed across these classes. Alternatively, if probability levels are used for the presence of commodity and forest classes, the use of precision and recall curves against probability level can help users determine an appropriate threshold suited to their needs (e.g. balanced precision or recall or maximize precision to minimize omissions for stratification purposes).

For exposure assessment, the most critical errors occur between commodities and forest, as these directly affect regulatory compliance and market access. For a map where the main class of interest is “commodities”, they are:

- commission errors, consisting in the points labelled in the reference as forest and mapped as the given commodity: potential undetected non-compliance (e.g. forest conversion incorrectly classified as coffee); and
- omission errors, consisting in the points labelled in the reference as the given commodity and mapped as forest: risk of unjust exclusion due to misclassification (e.g. deforestation false positive).

© nazmanm/Envato Elements

Validation reporting should explicitly quantify these error types rather than relying solely on overall accuracy or aggregate metrics.

Both field and high-resolution image annotation can be used for the accuracy assessment and validation, and can be adapted to different time periods (e.g. field data collection in 2025 for validating 2020 products). Historical image sources (e.g. Google Earth historical imagery), combined with local field knowledge, can be used to confirm that no major land-use change occurred between the mapped reference year and the validation observation.

Trained technicians carry out this process, following best practices to gather data from the field or to annotate an image as validating the modelled data product.

Capacity building to implement best practices for accuracy assessment should be integrated into mapping workflows and include all national partners, and produce transparent, replicable methods with publicly available validation data.

These assessments should be conducted at spatial scales and across strata that reflect local landscape realities, including smallholder agroforestry systems, heterogeneous mosaics and context-specific production practices.

Global products validation averaging results over large areas can mask systematic errors in regions where smallholders dominate. Local-level validation is therefore essential to guarantee that datasets are accurate where they matter most, in the landscapes where compliance decisions directly affect producers' market access and livelihoods.

In addition, local knowledge adds value to annotation processes. Evidence from a benchmarking exercise of 20 cocoa and forest mapping datasets in Ghana and Côte d'Ivoire (Reymondin ***et al.***, 2025) indicated that locally developed, country-level datasets produced by public institutions can achieve accuracy levels close to those of private service providers, while substantially outperforming global-scale datasets. Moreover, initiatives such as MapBiomas national collections or Global Pasture Watch illustrate the potential of national-scale mapping initiatives to incorporate locally grounded data, including of certain commodities, into large-scale monitoring and traceability systems, using datasets based on locally informed classification methods, regionally calibrated data and expert knowledge. This advantage is particularly strong when rigorous best practices for data collection, quality control and documentation are consistently applied at the local level.

To avoid overfitting and bias due to repeated evaluation, the final validation should be performed only once, on the finalized and frozen version of the map and model. During model development, a separate intermediate validation dataset should be used for feature engineering, model selection and iterative tuning.

Convergence of evidence

The CoE approach (d'Annunzio *et al*., 2024) is based on the principle that no single geospatial dataset can provide a complete picture of what has occurred on a given plot of land. By integrating datasets from multiple perspectives – such as crop data, tree cover loss and forest land-use information – CoE enables a more accurate assessment of land use and helps create a comprehensive understanding of what is most likely happening at plot level.

Within a CoE approach, maps of commodities can help disambiguate tree cover plots with agricultural use from tree cover plots with forest use, at a given point in time. This is achievable if data is made available; the next section looks into the requirements for sharing data and metadata.

Data sharing

Availability of end results

Commodity maps should be designed to support operational use, including supply chain risk assessment, land-use planning, monitoring and reporting. Public postprocessed maps (raster-based maps indicating presence or probability of commodities, validated through accuracy assessment) can be used by everyone for due diligence, audits and third-party checks. Such public products can also support land-use planning and activities like GHG reporting.

For instance, for large commercial operators, clear and public information on commodity presence at regulatory cutoff date under the European Union (EU) Deforestation Regulation (EUDR), 31 December 2020, can help EUDR due diligence to be done in a transparent way and reduce the administrative burden associated with transaction-based analyses. That in turn allows commercial operators to focus on identifying and addressing actual deforestation events.

For smallholders, public data can alleviate the risk of being excluded from markets requiring information related to commodity geolocation. In that sense, public data encourage inclusion rather than market exclusion, which is crucial for poverty reduction and long-term forest protection. This can also reduce the repetitive process many farmers face when asked to map and remap their farms to prove compliance under many different projects. Improved transparency and availability of geospatial information can also facilitate farmers' access to formal markets and financial services, including credit, by providing verifiable data that can support traceability, compliance and risk assessments used by buyers and financial institutions.

Open data can be challenged, corrected, updated and improved in a transparent manner. Initial investment in data production is an asset for further iterations. Therefore, public data on commodity presence represent a precompetitive asset for operators responsible for carrying due diligence on zero-deforestation production.

Public commodity maps can be visualized or downloaded through online platforms and repositories like GitHub, Africa GeoPortal, Hand-In-Hand, Source Cooperative, Zenodo, Dataverse or Google Earth Engine. Licensing for public data should allow the largest usage possible, including commercial purposes.

Availability of metadata

Metadata describing the methods used to obtain the data, including EO inputs, preprocessing approaches, metrics on reference data for both training and validation, and approaches for manual editions or corrections, but also limitations of the product and intended use cases, should accompany datasets to avoid misinterpretation and misuse.

That information can be shared through different pathways:

• minimum:	Publicly available methods document (e.g. whitepaper on GitHub)
• better:	Arxiv or other preprint
• ideal:	Peer-reviewed journal article

Beyond recommendations on publishing metadata, some good practices on data sharing modalities can be found below:

- Format: Open formats allow wide usage across platforms and tools and prevent vendor lock-in. Example: GeoTiffs for raster (or even including Cloud Optimized Tiffs, GeoTiffs with an internal organization that enables more efficient workflows on the cloud), GeoJSON for vector (for large vector datasets, compressed GeoPackage may be preferable, as its internal indexing improves data handling and query performance)
- Projection: WGS 84 (EPSG:4326) is a reasonable option for global data and is the only supported coordinate reference system for the GeoJSON specifications
- Versioning: Provides a way to track updates to a dataset and alert data users when there is a new version available

Availability of reference data

While publishing final map products is essential, sharing information about the underlying reference data used to train and validate those products is also important.

Access to information about reference data allows users to understand the sources of bias, representativeness and limitations embedded in a mapping system.

In many cases, global or national maps appear accurate overall but fail in specific contexts, especially in smallholder-dominated, agroforestry or heterogeneous tree crop landscapes, because the training data did not adequately represent these systems. Transparent information about the reference datasets enable independent evaluation, support reproducibility and encourage continuous improvement across the EO community.

A validation and benchmarking process that is locally grounded, transparent and stratified by relevant land-use types is critical to ensuring that EO-based solutions are both technically robust and socially equitable.

Where confidentiality constraints exist, anonymized, stratified or partially shared reference data (e.g. heat-map of reference data availability) should be made available to the greatest extent possible.

Towards a pool of reference data

Beyond the core recommendations of this paper, valid for single maps individually, it would benefit a large part of the global community to have an openly shared pool of anonymized, high-quality reference points to improve map accuracy and inclusiveness.

Several initiatives that follow the proposed good practices, in particular the Sample Earth initiative (Vantalon *et al.*, 2025), contribute to that potential pool of public data that can be brought together, within a CoE framework, to increase the transparency of supply chains at large.

Such a pool needs clear rules around data ownership, storage, update, access and reuse. Licensing ensures that data providers retain appropriate control over their contributions while enabling broader public benefits and to accelerate innovation. Data governance frameworks should define permissions for commercial use and sharing of field-collected geocoordinates.

UPSTREAM TO DOWNSTREAM TRANSFERABILITY OF DATA RIGHTS

The creation of accurate, locally relevant maps of forests, agriculture and agroforestry lands depends on a variety of inputs ingested from “upstream” data shared across public and private actors.

EO satellite missions by global space agencies, national public land and forest administrations, and rural cadastral land registries represent major public infrastructure investments without which EO maps of forest and commodity production would not be possible at scale. Upstream private sector data is equally integral to this digital ecosystem.

Farmers, land owners or managers, commercial timber harvesters, agricommodity small and medium-sized enterprises and supply chain operators accumulate significant digitized information on the land units they depend on, as do local service providers that may be supporting land managers with inputs, warehousing or other services.

As AI and machine learning have become central to EO mapping, models have come to rely on training data to be ingested along with the satellite-derived remote sensing data for a meaningful output (e.g. a commodity map) to be generated. As mentioned earlier, that output often requires additional ground data, very high-resolution imagery, or both for validation and quality assessment.

Whether the data originates in the private or public sector (upstream) and whether it is synthesized, analysed or used by the public or private sector (downstream), the value creation process is shared across the ecosystem of stakeholders, giving all actors a shared stake in the transferability of data and the right to create intellectual property (IP) from it.

The rest of this section provides recommendations on licensing pathways and informed consent aspects to move towards a public pool of reference data.

INTELLECTUAL PROPERTY CC-BY-4.0 AND TRANSFERABILITY OF DIGITAL RIGHTS

From an IP value creation perspective (maximizing public goods, private growth and share value creation) the best case would be to maximize the volume of public and private sector upstream data that have readily transferable Creative Commons Attribution 4.0 International License (CC-BY-4.0) IP rights accompanying its data from the time of its creation or collection and through all subsequent data sharing.

As most field data collection routinely entails consent to terms and conditions, these are a logical place to add optional or conditional consent from upstream stakeholders for CC-BY-4.0 IP creation data rights to be shared downstream and throughout the data chain of custody.

While transferable CC-BY-4.0 data rights should be integrated to upstream data collection systems to maximize shared value creation downstream, the acknowledged intention that data will be shared, exchanged and used beyond a single limited purpose requires consent terms and conditions to also address the personal, community and commercial risks and opportunities associated with third-party data use.

DATA PROTECTION LAWS ACCORDING TO LOCAL JURISDICTION

All stakeholders are situated within data governance laws of varying complexity and maturity according to national jurisdictions with varying approaches to personal data protection, government data access rights, data sovereignty, IP licensing and national data sovereignty laws that may regulate transfer of data abroad.

Adherence to national data protection and privacy laws (see "Data protection laws" section in the References for examples) should be actively addressed as part of continuous improvements to consent terms and conditions.

A great deal of data protection focuses on reducing risks to individuals by minimizing the collection and retention of PII. In accordance with this, the upstream data collection consent practices should:

- minimize paper records in favour of digital systems;
- avoid collection of information on names, dates of birth, gender or other easily cross-referenced PII;
- implement timely disposal or destruction of any personal information; and
- employ data anonymization methods as far as possible, such as use of randomized unique identifiers in lieu of national ID numbers or similar

ANONYMOUS PERSISTENT IDENTIFIERS

When persistent identifiers are required, a recommended approach is to compute them as follows:

hash(owner_id + owner_secret + salt). "owner_secret" allows the data owner to later prove ownership and request deletion of their data if needed.

Combining the secret with a "salt" helps reduce the risk of dictionary attacks, particularly when "owner_id" comes from a small or predictable set of values. Hashing also helps limit direct fingerprinting and reduces the ability to link the same individual across independent datasets. For example, in Reymondin ***et al.*** (2024) you can see an example of use of hash to anonymize ID (in this case smartphone MAC addrees).

In cases where the user explicitly consents, these constraints may be relaxed in controlled environments to allow deeper analysis. However, practitioners should still account for spatial re-identification risks. For example, strong spatial correlations, such as a farm located next to a large, well-known commodity plot, may allow individuals to be re-identified even when identifiers themselves are hashed.

LEVELS OF CONSENT

While there is a great deal of emphasis on privacy and data laws, it is also worth acknowledging and considering that data sharing also has positive commercial and public value in many scenarios. The act of digitizing information is done for a stated benefit with consent given out of direct or public interest. Informed consent means that those participating in data collection agreements should be informed of the risks as well as the benefits. There are many cases where sharing information has direct value, such as access to credit or markets, as well as many circumstances where it can present a business or security risk, such as increased credit costs, unequal application of tax laws or compromised personal safety should the data be leaked.

As a general practice, informed consent should be addressed using the many guidance documents available. In the context of land use and mapping tree crop commodities, however, it is important to recognize that a general “informed consent” may be adequate, but should be distinguished from the term “free, prior and informed consent” (FPIC), which has arisen in relation to frameworks intersecting Indigenous or recognized communities where collective, extralegal or other land rights may apply beyond the standard informed consent.

A best-case digitized terms and conditions would be one that integrates these issues of transferable CC-BY-4.0 IP rights, adherence to local data laws, informed consent (and FPIC) to risks and benefits and minimization of PII. More importantly, these digital rights must be captured and transferred across data systems in attachment to the data which serves to drive value creation in the first place: geometries, minimum attributes and metadata required to train, validate and quality control or quality assure EO-based mapping of forests and tree crop commodities.

Conclusion

The mapping of tree crop commodities represents both a critical opportunity and a complex challenge in the global effort to address deforestation and promote sustainable supply chains. As regulatory, market and societal expectations continue to rise, the need for reliable, transparent and comparable geospatial data has become increasingly urgent. However, the structural complexity of agroforestry systems, limitations of EO technologies, and gaps in reference data and governance frameworks hinder the production of consistent and accurate maps.

These forest monitoring good practices on tree crop commodities provide a pragmatic and flexible foundation to overcome such challenges by promoting shared definitions, methodological alignment and best practices in data collection, modelling, validation and data sharing. Importantly, they emphasize that technological solutions alone are insufficient; progress also depends on improved data governance, collaboration and trust among stakeholders.

By encouraging transparency, interoperability and the inclusion of local knowledge and smallholder realities, this document support more equitable and effective decision-making. Ultimately, their adoption can strengthen the credibility of deforestation monitoring systems, reduce compliance burdens and enable more responsible sourcing practices.

References

This section provides a non-exhaustive selection of published references relevant to the topic of mapping tree crop commodities.

Accountability Framework Initiative. 2021. Traceability.

D'Annunzio, R., O'Brien, V., Arnell, A., Neeff, T., Fontanarosa, R., Valbuena Perez, P., Shapiro, A.C., Sanchez-Paus Díaz, A., Merle, C., Vega, J. & Fox, J. 2024. Towards a digital public infrastructure for deforestation-related trade regulations – What is in that plot? (Whisp) solution to implement convergence of evidence. Rome, FAO.

Berger, K., Foerster, S., Szantoi, Z., Hostert, P., Foerster, M., Van De Kerchove, R., Vancutsem, C., Schweitzer, C., Masolele, R., Reiche, J. and Dowell, M., 2025. Evolving Earth observation capabilities for recent land-related EU policies. *Land Use Policy*, 158: 107749.

Berger, K., Herold, M. & Szantoi, Z. 2025. Earth observation as enabler for implementing the EU regulation on deforestation-free products. *npj Climate Action*, 4: 68.

BNETD. 2024. Côte d'Ivoire 2020 land-cover map.

Branthomme, A., Merle, C., Kindgard, A., Lourenço, A., Ng, W.-T., D'Annunzio, R. & Shapiro, A. 2023. How much do large-scale and small-scale farming contribute to global deforestation? Results from a remote sensing pilot approach. Rome, FAO.

Reymondin, L., Vantalon, T., Luong, P., Jairo Tello, J., Perez, J. Matarasso, M., Howard, E. 2025. Establishing best practices and evaluating data and methods for forest monitoring

Clinton, N., Vollrath, A., d'Annunzio, R., Liu, D., Glick, H.B., Descals, A., Sullivan, A., Guinan, O., Abramowitz, J., Stolle, F. &Goodman, C., 2024. A community palm model. *arXiv preprint arXiv:2405.09530*.

Degerickx, J. 2026. From existing crop maps to training samples: Leveraging local crop-mapping initiatives to strengthen WorldCereal's global crop mapping. In: *VITO Remote Sensing*. Mol, Belgium.

Descals, A., Gaveau, D.L., Wich, S., Szantoi, Z. &Meijaard, E., 2024. Global mapping of oil palm planting year from 1990 to 2021. *Earth System Science Data*, 16(11): 5111–5129.

Di Gregorio, A. & Jansen, L. J. M. 2000. Land Cover Classification System (LCCS): Classification Concepts and User Manual. Fao 53, 179. https://www.fao.org/4/x0596e/X0596e01e.htm

Escobar-López, A., Castillo-Santiago, M.Á., Mas, J.F., Hernández-Stefanoni, J.L. &López-Martínez, J.O. 2024. Identification of coffee agroforestry systems using remote sensing data: a review of methods and sensor data *Geocarto International*, 39(1): 2297555.

European Commission. 2025. Traceability and geolocation of commodities subject to EUDR – Green Forum.

FAO. 2023. Terms and Definitions – FRA 2025. Forest Resources Assessment (FRA) Working Paper 194. Rome, Italy

FAO & ISO. 2025. ISO19144-1: Geographic information – classification systems: Classification system structure. Part 1: ISO 19144-1:2009 with Cor.1:2012 (Edition 1). Geneva, Switzerland, ISO.

FAO & ISO. 2025. ISO19144-2: Geographic information – classification systems—Part 2: Land cover meta language (LCML). Part 2: ISO 19144-2:2023 (Edition 2). Geneva, Switzerland, ISO.

FAO & ISO. 2025. ISO19144-3: Geographic information – classification systems: Land use meta language (LUML). Part 3: ISO 19144-3:2024 (Edition 1). Geneva, Switzerland, ISO.

Forest Data Partnership. 2025. Commodity recognition book.

Freitas Beyer, J., Köthke, M., & Lippe, M. 2025. Assessing the suitability of available global forest maps as reference tools for EUDR-compliant deforestation monitoring, *Remote Sensing*, *17*(17): 3012.

Freitas Beyer, J., Köthke, M., & Lippe, M. 2026. GEOS-EUDR: Forest agreement layer (FAL) – Google Earth engine script (Version 1). Zenodo.

HCV Network. 2022. Common guidance for the management and monitoring of high conservation values.

Hunt, D. A., Tabor, K., Hewson, J. H., Wood, M. A., Reymondin, L., Koenig, K., Schmitt-Harsh, M., & Follett, F. 2020. Review of remote sensing methods to map coffee production systems. *Remote Sensing*, 12(12): 2041.

ISEAL. 2020. Guidance on the collection of polygon location data.

JAXA (Japan Aerospace Exploration Agency). 2003. Earth Observation Satellite and Mechanism of Observation. [Cited 10 January 2026] https://www.satnavi.jaxa.jp/en/satellite-knowledge/whats-eosatellite/observation/index.htm

Jonckheere, I., Hamilton, R., Michel, J.M. & Donegan, E., eds. 2024. *Good practices in sample-based area estimation. White Paper*. Rome, FAO.

Kalischek, N., Lang, N., Renier, C., Daudt, R.C., Addoah, T., Thompson, W., Blaser-Hart, W.J., Garrett, R., Schindler, K. &Wegner, J.D. 2023. Cocoa plantations are associated with deforestation in Côte d'Ivoire and Ghana. *Nature Food*, 4(5): 384–393.

Kerner, H., Chaudhari, S., Ghosh, A., Robinson, C., Ahmad, A., Choi, E., Jacobs, N., Holmes, C., Mohr, M., Dodhia, R. &Ferres, J.M.L. 2025. Fields of the world: A machine learning benchmark dataset for global agricultural field boundary segmentation. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27): 28151–28159.

Maastricht University. 2023. Smallholder-oriented data governance principles.

Maskell, G., Chemura, A., Nguyen, H., Gornott, C. &Mondal, P., 2021. Integration of Sentinel optical and radar data for mapping smallholder coffee production systems in Vietnam. *Remote Sensing of Environment*, 266: 112709.

Meridia. 2024–2025. EUDR FAQ & guidance. https://www.meridia.land/eudr

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E. and Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change . *Remote Sensing of Environment*, 148: 42–57.

Orlowski, K., Meroni, M., De Marzo, T., Imala, V., Ayanga, J. et al. 2026. Guidelines for harmonised VHR-based data collection for GeoAI applications in perennial crop mapping for historical reference years. Luxembourg, Publications Office of the European Union.

Otekunrin, O.A. 2025. Mapping cocoa research (Theobroma cacao L.) in Africa: Production, trade outlook, and scientometric insights. *Discover Agriculture, 3*(1):111.

Parreiras, T.C., Santos, C.D.O., Bolfe, É.L., Sano, E.E., Leandro, V.B.S., Bayma, G., Silva, L.A.P.D., Furuya, D.E.G., Romani, L.A.S. &Morton, D. 2025. Dense time series of harmonized Landsat Sentinel-2 and ensemble machine learning to map coffee production stages. *Remote Sensing, 17*(18): 3168.

Preferred by Nature. 2025. Identify good practices on data governance for production plot information in commodity supply chains.

PROFOREST. 2025. Open data for supply chain due diligence.

Reymondin, L., Vantalon, T., Luong, P., Tello, J. J., Perez, J., Matarasso, M., and Howard, E. 2025. Establishing best practices and evaluating data and methods for forest monitoring.

Reymondin, L., Vantalon, T., Pham, H.T. et al. 2024. Using free Wi-Fi to assess impact of COVID-19 pandemic on traditional wet markets in Hanoi. Food Sec. 16, 223–241 https://doi.org/10.1007/s12571-023-01417-w

Stehman, S.V. 2014. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *International Journal of Remote Sensing*, 35(13): 4923–4939.

Stehman, S.V., & Foody, G.M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, 231: 111199.

Tyukavina, A., Stehman, S.V., Foody, G.M., Bontemps, S., See, L., Olofsson, P., Tsendbazar, N.E., Radoux, J., Komarova, A., Serre, B.M. &de Bruin, S., 2025. *Land cover and change map accuracy assessment and area estimation good practices protocol* v1.1. Brussels, Committee on Earth Observation Satellites.

Vantalon, T., Luong, P.T., Perez Escobar, J.A., Tello Dagua, J.J., Phan, T.V., Nguyen, H. &Reymondin, L. 2025. Sample Earth: Machine-learning-ready land-cover reference Dataset.

Van Tricht, K., Degerickx, J., Gilliams, S., Zanaga, D., Battude, M., Grosu, A., Brombacher, J., Lesiv, M., Bayas, J.C.L., Karanam, S. &Fritz, S. 2023. WorldCereal: A dynamic open-source system for global-scale, seasonal, and reproducible crop and irrigation mapping. *Earth System Science Data, 15*(12): 5491–5515.

Wood, M.A., Gauttier, F., Bellfield, H., West, C. &Indenbaum, R.A. 2021. Data requirements to tackle global deforestation through mandatory due diligence.

WRI (World Resources Institute). 2022. Data& Methods. Key Terms and Definitions. New York, USA, UN. [Cited 10 January 2026] https://gfr.wri.org/key-terms-definitions

Yue, Z. &Xiao, C., 2025. A meta-review of remote sensing for rubber plantations. *International Journal of Applied Earth Observation and Geoinformation*, 141: 104625.

Zhang, Y., Woodcock, C.E., Arévalo, P., Olofsson, P., Tang, X., Stanimirova, R., Bullock, E., Tarrio, K.R., Zhu, Z. &Friedl, M.A. 2022. A global analysis of the spatial and temporal variability of usable Landsat observations at the pixel scale. *Frontiers in Remote Sensing, 3*: 894618.

Data protection laws

The below table provides some data protection law examples from countries that are key exporters and importers of tree crop commodities.

Jurisdiction	Data Protection Law(s)	Official URL(s)
Indonesia	UU No. 27/2022 – Pelindungan Data Pribadi (PDP Law)	https://peraturan.go.id/.../uu-no-27-2022
Brazil	Lei 13.709/2018 – Lei Geral de Proteção de Dados (LGPD)	https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm
Nigeria	Nigeria Data Protection Act (NDPA)	https://ndpc.gov.ng
European Union	Regulation (EU) 2016/679 – General Data Protection Regulation (GDPR)	https://eur-lex.europa.eu/eli/reg/2016/679/oj
Colombia	Ley 1581 de 2012 + Decreto 1377 de 2013	www.funcionpublica.gov.co/eva/gestornormativo
Ghana	Data Protection Act, 2012 (Act 843)	https://www.dataprotection.org.gh
Côte d'Ivoire	Loi N° 2013-450 du 19 juin 2013 Relative à la protection des données à caractère personnel	https://www.artci.ci

Existing initiatives

This section provided several examples of initiatives and public or private partnerships that contribute to generating relevant and up-to-date commodity maps.

Those initiatives constitute, at least partially, positive examples of compliance with the current good practices.

Forest Data Partnership

Consortia: FAO, Google, WRI, SIG, UNILEVER, LDC, CIAT, SERVIR
Scope: Global
Commodities: palm, cocoa, coffee, rubber

https://developers.google.com/earth-engine/datasets/publisher/forestdatapartnership

The Forest Data Partnership strengthens collaboration and application around global monitoring of commodity-driven deforestation, forest degradation and restoration efforts across the globe. Governments and companies around the world are pledging to help end deforestation and accelerate restoration to avert the worst impacts of climate change, protect against biodiversity loss and safeguard the many benefits of forests to people and nature.

The partnership aims to halt and reverse forest loss from commodity production by collaboratively improving global monitoring and supply chain tracking, as well as accelerating restoration. It aligns partners around the data and ensures access for stakeholders across sectors to open-source data on geospatial forest-risk commodities that is consistent and validated. The result is credible, systematic monitoring, verification and accountability towards progress in reducing commodity-driven deforestation and restoring degraded lands.

Sample Earth

Consortia: CIAT, Satelligence, WCF & Forest Data Partnership
Scope: Vietnam, Ghana
Commodity: Cocoa, coffee

https://doi.org/10.7910/DVN/U7HWY1

Sample Earth, an international initiative led by the Alliance of Bioversity International and CIAT, aims to improve the accuracy of existing deforestation-free compliance tools by ensuring they correctly represent smallholder farmers, whose land is often misclassified as forest. The initiative does this by generating and openly sharing hundreds of thousands of accurately labelled GPS points in smallholder landscapes so that compliance tool developers can build fairer, more reliable assessments. More accurate maps will be vital to millions of smallholders, consumers and businesses, who will soon need to comply with impending deforestation-free regulations set by the European Union. Increased collaboration is critical to expand Sample Earth's mapping system, which is designed to ease proof of compliance.

The first version of the Sample Earth reference dataset includes about 100k time-stamped geolocation data from coffee and cocoa farms, as well as 9 other land-cover classes such as oil palm, timber, rubber, and 88 subclasses, including complex systems like coffee intercropped with black pepper. The dataset focuses on two key regions: Vietnam's Central Highlands, the world's second-largest coffee-producing country, and Ghana, the world's second-largest cocoa producer.

In addition to providing the tools (better training datasets) to build more accurate maps, the Sample Earth team aims to create a benchmark for map quality and inclusiveness. This includes a built-in map-improvement mechanism, where mapmakers can access confidential land-use data to evaluate, and improve, their maps.

COFFEE CANOPY PARTNERSHIP

Consortia:	JDE Peet's, JDE Peet's, Louis Dreyfus Company, Sucden, Neumann Kaffee Gruppe, Touton, Sucafina, Tchibo, Airbus, supported by the UK Foreign, Commonwealth & Development Office (FCDO) and in technical collaboration with FAO.
Scope:	Burundi, Rwanda, Uganda, Kenya, Tanzania, Ethiopia

This initiative aims to develop open-source, robust and unique 2020 and 2025 maps of all coffee plots, a first critical step towards strengthening traceability, supporting compliance with sustainability requirements, and supporting farmers' access to EU markets.

By accurately identifying coffee cultivation areas using satellite imagery and AI, the project will enable remediation and avoidance of deforestation post 2020; reduce the misclassification of coffee agroforestry as forest and consequently help farmers retain access to key markets, particularly under the EUDR; address root causes of deforestation, such as poverty, by identifying at-risk areas in need of support; contribute to positive origin image with consumers by assuring no coffee is produced on deforested land. These maps, starting from the EUDR cutoff date of December 2020, will offer actionable, transparent data to drive more equitable and sustainable practices across the global coffee sector.

SUSTAINABLE COCOA INITIATIVE (SCI)

Consortia:	EU, GIZ, FAO, EFI, JRC
Scope:	Côte d'Ivoire, Ghana, Cameroon
Commodity:	Cocoa

https://zerodeforestationhub.eu/projects/sci/

The Sustainable Cocoa Initiative promotes a sustainable, equitable and productive cocoa global value chain, linking producing countries in West Africa with the European consumer market and broader global cocoa sector. It addresses challenges like deforestation, child labour and living income. The initiative supports the implementation of the Sustainable Cocoa Alliance roadmap – adopted by Côte d'Ivoire, Ghana and Cameroon – and includes measures to develop baseline forest and land-use and commodity maps to promote transparency and publicly accessible data on forests and land cover.

This includes the development by the Bureau National d'Etudes Techniques et de Développement of a freely available national land-cover map of Côte d'Ivoire for 2020, which the Joint Research Centre (JRC) independently validated and which will be updated in 2026.

The FAO Investment Centre supported the development of 2020 tree-cover and land-cover data with a multistakeholder national technical working group in Cameroon. A cocoa probability map is under development with the technical and financial support of the EU and FAO. This initiative is using innovative satellite technology and AI foundation models to map agroforestry and open canopy cocoa, its expansion and its impact on forests. An independent validation of the cocoa probability model map is being conducted by the JRC through the activation of the Copernicus4GEOGLAM map validation service of the Copernicus Land Monitoring Service , whose objective is to support mapping and accuracy assessments in the context of the EUDR. Within this framework, the validation service adds value by providing an independent, objective and transparent accuracy assessment of these products. This independent evaluation is designed to build trust and confidence in the resulting maps and to enhance their credibility, acceptance and uptake among national and international stakeholders.

In Ghana, the European Forest Institute (EFI) and FAO (through the Aim4Forest project) are supporting the Forestry Commission to produce a 2020 forest and non-forest map for the cocoa region to support the Ghana Cocoa Board and operators' deforestation risk analysis.

TECHNICAL FACILITY ON DEFORESTATION-FREE VALUE CHAINS

Consortia:	EFI
Scope:	Global
Commodity:	Coffee, cocoa, rubber, palm oil

https://efi.int/partnerships/valuechains

Under the Team Europe Initiative, the Technical Facility on Deforestation-Free Value Chains (TF) is a global instrument supporting producing countries in preparing for EUDR implementation.

To address the commodity-specific nature of deforestation risk, the TF supports the development and validation of crop and agroforestry maps that complement forest data, particularly in smallholder-dominated landscapes. The TF's approach prioritizes building on existing national forest and land-use monitoring systems, ensuring they are fit for both domestic forest governance and EUDR due-diligence needs. Across partner countries, the TF has assessed the availability, quality and usability of national forest and land-cover datasets, supported the development or upgrading of national maps, and facilitated inter-agency coordination between forest authorities, commodity agencies and land administrations.

Regarding tailored support: In Ethiopia, the TF initiated upgrades of the national forest map alongside coffee-relevant spatial datasets; in Kenya, it developed a national coffee probability map, including agroforestry considerations, in collaboration with national agencies and the JRC; in Tanzania, it assessed national forest and land-cover datasets linked to EUDR preparedness; in Sierra Leone and Liberia, it is supporting the updated national forest maps to support deforestation risk analysis in cocoa landscapes; and in Honduras, it supported the validation of national coffee and land-cover datasets with national authorities and the JRC to improve their fitness for EUDR-related deforestation checks. These country results have been complemented by methodological work with the JRC to establish common criteria for dataset accuracy and fitness-for-purpose by targeted analytical support and policy dialogue in countries such as Uganda, Sri Lanka and Lao People's Democratic Republic, ensuring that deforestation risk analysis is applied in a risk-based manner aligned with national systems and EUDR requirements

WORLDAGROCOMMODITIES

Consortia:	ESA, GAF-AG, VITO, Wageningen Uni, GFZ, TerraSphere, Meridia, e-Geos
Scope:	Global
Commodity:	Coffee, cocoa, rubber, palm oil

https://agrocommodities.info/

The primary focus of the ESA-led WorldAgroCommodities project is on the development of a preoperational monitoring system to support the implementation of the EUDR by EU Member States to support the checks to be made as part of the Due Diligence Statement inspection process.

This system will align with the requirements and needs of the competent national authorities (CNAs) responsible for monitoring the compliance of operators and traders. A comprehensive consultative process with CNAs was initiated at the project onset to gather detailed requirements and will continue throughout the duration of the project.

The user requirement phase identified two main steps of the inspection process on which to focus: i) a fully automated tool to sieve through Due Diligence Statements to identify those requiring more detailed inspection through the identification of potential non-compliance (e.g. deforestation post 2020 or inconsistencies with declared commodity); ii) inspection-level work to support the identification of non-conformance.

Initial tests were carried out on representative test sites around the world based on priority countries identified from CNAs with a set of criteria aimed at ensuring a representative sample of test sites across different production systems and regions. Ten sites of a minimum area of 100 km² (identified from H3 grid cells) were selected, and several methods to identify deforested areas and commodity types were tested through a benchmarking approach.

The preliminary design of the system is based on a dynamic mapping approach in which the CNA inspector will be able to run a ML/DL model on the fly to identify deforested and commodity areas. The system will adopt a cloud-based platform-agnostic architecture to allow the integration within the CNA's own system. The outcome of the benchmarking process is now being implemented on OpenEO/CDSE. The next steps will be to apply the tool over larger geographical areas covering at least four countries for each commodity, validate the results and develop a series of use cases in collaboration with the CNAs.

Digital Integration of Agricultural Supply Chains Alliance (DIASCA)

Consortium: GIZ, FAO, COSA, CGIAR, ITC, and many more
Scope: Global
Commodity: multiple

www.sustainable-supply-chains.org/topics/digitalisation-traceability/diasca/

The DIASCA is a neutral, precompetitive, and inclusive multistakeholder network that promotes digital public infrastructure (DPI) in agricultural supply chains. The goal is to strengthen digital integration, interoperability and inclusive participation to enable sustainable, transparent and efficient agricultural supply chains worldwide. DIASCA fosters DPI in agricultural supply chains to support smallholder farmers, protect natural ecosystems and make supply chain due diligence more accessible and efficient for all actors. Among others, a core component of DIASCA's work is the integration of geospatial data. This includes farm-level data, satellite imagery and land-use mapping to improve traceability and transparency. The DIASCA is steered by Sustainable Agricultural Supply Chains Initiative (SASI), implemented by Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) on behalf of the German Federal Ministry for Economic Cooperation and Development .

GEOS-EUDR

Consortium: GIZ SAFE and GIZ SASI, Nuup A.C., OroVerde Tropical Forest Foundation, IPB University Indonesia, Olam Agri Indonesia
Scope: Global with pilot studies in Brazil, Indonesia, Mexico
Commodity: Multiple; Coffee, Cacao, Rubber in case of pilot studies

www.sustainable-supply-chains.org/topics/digitalisation-traceability/diasca/

The GEOSpatial remote sensing–based detection of land-use change in support of the EU regulation on deforestation-free supply chains (EUDR) project (GEOS-EUDR) accompanies the implementation of the EU regulation on deforestation-free supply chains (EUDR) from a scientific perspective. It aims to contribute to the operational feasibility of the regulation with a focus on geolocation and the detection of deforestation and forest degradation using available remote sensing–based products and geocomputation. GEOS-EUDR considers the feasibility of an EUDR-compliant implementation along the entire stakeholder chain, from geodata collection by producers (often organized in smallholder structures and farming cooperatives) to the desk-based monitoring by national authorities. The project shows possibilities and limitations in differentiating between land-use types, with a particular focus on tree-based systems such as agroforestry and timber plantations, and thus provides an important input for global policymakers including EU COM, national government agencies and operators. The project is analysing the implementation of the EUDR using global reference maps and local reference information in three pilot regions. Possibilities and limitations in differentiating different land-use types for the context of the pilot areas will be determined, as well as findings on the interpretation of deforestation control and its jurisdiction. Desktop-based results will be compared with in situ data in case of the pilot studies to assess the accuracy of the remote sensing–based approaches. The project co-operates with diverse country partners for the purpose of the pilot study regions.

WITH THE FINANCIAL SUPPORT OF:

UK Government

Global Gateway

Co-funded by the European Union

Team Europe Initiative on Deforestation-free Value Chains

german cooperation
DEUTSCHE ZUSAMMENARBEIT

Ministry of Foreign Affairs

Implemented by
giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

ISBN 978-92-5-140699-1

9 789251 406991

CD9622EN/1/05.26